PRM97

DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF A CANINE DERMATITIS QUALITY OF LIFE QUESTIONNAIRE: RESULTS FROM THREE CLINICAL TRIALS

Postorino-Reeves N[1], Brohan E[2], Tatlock S[2], Panter C[2], Amodie D[1], Gasper S[1]
[1]Zoetis, Florham Park, NJ, USA, [2]Adelphi Values Ltd, Manchester, UK

OBJECTIVES: Atopic dermatitis is a common allergic skin disease in canines, characterized by itching and scratching potentially leading to hair loss, excoriations and scabbing. Atopy can substantially impact both the owner and canine's quality of life (QoL). The objective was to develop a robust, pet-owner completed, questionnaire to assess QoL in canines with atopic dermatitis and their owners. **METHODS:** A 34-item draft questionnaire was developed through review of the literature and clinical input. This measure was included in three clinical trials of APOQUEL: a single arm, open-label continuation study in the United States (n=247); and two randomized, controlled, blinded clinical trial in Australia (n=124; n=226). Questionnaire data from each trial was individually subject to item response, dimensionality, and scale-level analyses. Mean change in QoL scores were also analyzed as an indicator of treatment efficacy. **RESULTS:** A total of 597 dog owners completed the questionnaire. Quality of completion was high (maximum 4.9% missing data). Across all studies, seven items consistently displayed ceiling effects (range: 23.5%-86.3%) and four displayed floor effects (range: 21.6%-70.4%). These results, together with the item discrimination index, confirmatory and exploratory factor analyses (CFA, EFA), item-total correlations and internal consistency, were used to guide item reduction. This resulted in a final measure on which known-groups validity and responsiveness to change were examined. Overall, owners of dogs on APOQUEL demonstrated a positive mean change on six items across all three datasets, indicating improvement in these aspects of QoL. **CONCLUSIONS:** This study describes the development and psychometric evaluation of an innovative measure of QoL in canines with atopic dermatitis and their owners. Findings from each study, along with consideration of the clinical relevance were used to support item reduction decisions. The final questionnaire then underwent scale-level psychometric evaluation. Consistent improvements in QoL were seen across all three studies.

PRM98

DATA QUALITY WITH MIXED-MODE ADMINISTRATION OF THE SHORT FORM-36

Broering JM[1], Paciorek A[1], Carroll P[2], Wilson LS[1], Litwin MS[3], Miaskowski C[1]
[1]University of California San Francisco, San Francisco, CA, USA, [2]University of California, San Francisco, San Francisco, CA, USA, [3]University of California Los Angeles, Los Angeles, CA, USA

OBJECTIVES: To examine the effects of survey mode of administration on data quality when using a mixed-mode (i.e., paper-mode vs. web-mode) approach. **METHODS:** A cross-sectional wave of 4,836 participants active in the Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE) registry, self-selected to complete the Medical Outcomes Short-Form 36 instrument. Data quality parameters compared response rate; data completeness at the item and computable scale score; differences in mean scale scores; and range of values (i.e., ceiling and floor effects) by mode. **RESULTS:** The majority of participants, 4,376 (90%), opted for paper-mode. Overall response rate was 77%; paper-mode 76% versus 88% web-mode (p <.01). Paper-mode respondents were older, 70.7 (SD 8.8) versus 66.8 (SD 8.5) (p <.001). Paper-mode participants had significantly more missing items (mean .67 vs. .10, p <.0001) and significantly lower proportion of computable scale scores for bodily pain, general health, vitality, social functioning, and mental health (p < .03). The social functioning scale had the largest difference for the proportion with missing data, 92.3% vs. 99.5% (p = .00). Mean scale scores were lower among paper-mode participants for all eight scales and the physical and mental component summary scores (p <.03). Significantly lower mean scale scores with small differences in effect size were noted for paper versus web-mode for the physical component summary, 47.5 vs. 50.2 (D = .30); physical functioning, 80.2 vs. 88.2 (D = .36); and vitality, 64.8 vs. 71.2 (D = .32). Four scales demonstrated a full range of values (0 to 100). For the web-mode, four scales did not achieve a floor value: Vitality (10 – 100); mental health (28 – 100); bodily pain 23 – 100; and general health (15-100). **CONCLUSIONS:** We observed significant differences in data quality with mixed-mode survey administration but effect-size differences were small suggesting that mixed-mode administration did not introduce significant measurement differences.

PRM99

ALGORITHMS TO ESTIMATE HEALTH UTILITIES FROM TOTAL JOINT ARTHROPLASTY DISEASE-SPECIFIC MEASURES

Odum SM[1], Troyer JL[2]
[1]UNC-Charlotte, Charlotte, NC, USA, [2]UNC-Charlotte, Charlotte, NC

OBJECTIVES: Orthopedic researchers studying total joint arthroplasty (TJA) often use disease-specific measures (DSMs) of patient health as cost-effectiveness outcomes. However, for cost-utility analysis, health utility scores are needed. The study objective was to develop regression algorithms to map five common disease specific TJA outcome measures to three preference-based utility scores. **METHODS:** An online survey was completed by 438 total hip arthroplasty (THA) patients and 550 total knee arthroplasty (TKA) patients. THA patients completed the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC®), Harris Hip Score (HHS), and the Hip Disability and Osteoarthritis Outcomes Score (HOOS). Knee patients completed the WOMAC®, Knee Society Score (KSS), and Knee Disability and Osteoarthritis Outcomes Score (KOOS). All patients completed three preference based questionnaires, the SF-6D, EQ-5D and HUI-3, and responses were used to calculate health utilities. A total of 30 THA and 30 TKA mapping models were estimated for each possible pairing of utilities and DSMs, and prediction performance was considered to pick the best DSM/utility match. **RESULTS:** For THA, the regression model with HOOS subscores most precisely estimated an EQ-5D utility. For the HOOS/EQ-5D model, the MAE was .05, the RMSE was .06 and the ME was nearly zero. The best performing TKA model mapped the KSS to the EQ-5D. The MAE was .05, the RMSE was .06 and the ME also approximated 0.00 for the KSS/EQ-5D. These prediction errors are small which indicates that these models can accurately estimate the EQ-5D utilities. However, they should only be used in conjunction with average patient characteristics and not with individual level data. **CONCLUSIONS:** Clinicians and researchers can input their disease specific data into these models to estimate health utilities to consider the cost-effectiveness of osteoarthritis-related interventions relative to interventions for very different diseases and conditions.

PRM100

EQUIVALENCE OF ELECTRONIC AND PAPER ADMINISTRATION OF PATIENT REPORTED OUTCOME MEASURES: A SYSTEMATIC REVIEW AND META-ANALYSIS

Wild D[1], Muehlhausen W[1], O Donohoe P[2], Doll H[1]
[1]ICON, Oxford, UK, [2]CRF Health, London, UK

OBJECTIVES: To conduct a systematic review and meta-analysis of the equivalence between electronic and paper administration of patient reported outcome measures. **METHODS:** A systematic literature review was undertaken in which 1,997 records were identified from which 73 studies met our inclusion criteria. The data were subject to a random effects meta-analysis. Pooled estimates of correlation and mean difference were estimated. The modifying effect of particular study characteristics was explored by calculation of pooled values for studies grouped by: mode of administration, year of publication, study design, and time interval between administrations. **RESULTS:** 435 individual correlations were extracted across all publications. The random effects pooled correlation coefficient was 0.875 (95% CI 0.866 to 0.883). Correlations were available for 56 studies, with average values still highly variable (I^2=91.56). After excluding 20 studies with outlying values the I2 was 56.63, with an overall random effects pooled correlation coefficient of 0.870 (95% CI 0.856 to 0.882). In terms of factors that might explain the heterogeneity, there was a statistically significant difference in pooled correlation estimates between years, with a tendency for agreement to be greater in more recent studies (fixed p<0.001, random p<0.001). Paper vs hand-held agreement was significantly higher than paper vs web agreement (0.894 vs 0.855), with non-overlapping 95% CI. **CONCLUSIONS:** The present study supports the conclusion of previous a previous manuscript (Gwaltney et al 2008) showing that written assessments administered on paper are the same as written assessments on an electronic device. This study has confirmed this conclusion and has also broadened the comparison out to include IVRS assessments and breaking the comparison down by hand-held vs internet administration. The results suggest that scores obtained by different modalities are directly comparable.

PRM101

DEVELOPMENT OF THE BEHAVIOR RATING INVENTORY OF EXECUTIVE FUNCTION (BRIEF) IN FIVE LANGUAGES

Vasarri S[1], Acquadro C[2], Isquith PK[3]
[1]Mapi, Lyon, France, [2]Mapi Research Trust, Lyon, France, [3]Dartmouth University, Norwich, VT, USA

OBJECTIVES: The Behavior Rating Inventory of Executive Function (BRIEF) is the most widely used questionnaire developed for parents and teachers of school-age children to assess executive function behaviors of children and adolescents in the school and home environments. It is composed of 86 items organized in eight clinical scales (Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, Monitor) and two validity scales (Inconsistency and Negativity). It is used to evaluate children/adolescents (5-18 years old) with a wide spectrum of developmental and acquired neurological conditions. The objective of this study was to develop the BRIEF in four Balto-Slavic languages (Bulgarian, Latvian, Lithuanian and Serbian) and one Uralic language (Estonian). **METHODS:** The following methodology was used: (1) Clarification of concepts with the developers; (2) Forward/backward step: two forward translations by native translators, reconciliation of the forward translations; one back-translation by an English-speaker fluent in the target language, and review by the developer. **RESULTS:** The translation process did not reveal any cultural issues since most of the concepts assessed were cross-culturally relevant. The main difficulties consisted in finding conceptual equivalents of the original items with strong idiomatic content. For instance, the most problematic items were items 18, 55 and 71. Item 71 [Lies around the house a lot ("couch potato")] raised difficulties because of the expression "couch potato" which had to be removed in most of the translations. The intended meaning (the child loafs around the house, lying on the couch, doing nothing productive) was rendered in all target versions. Deletion of this expression in the original US version was also discussed with the developers. **CONCLUSIONS:** The cross-cultural adaptation of the BRIEF into Balto-Slavic and Uralic languages required an international collaboration. The involvement of the developers during the whole process enabled the production of conceptually equivalent and culturally appropriate measures.

PRM103

ASSESSING THE CONTENT VALIDITY OF THE ENDOMETRIOSIS SYMPTOM DIARY AND ENDOMETRIOSIS IMPACT SCALE

Gater A[1], Wichmann K[2], Seitz C[2], Gerlinger C[2], Taylor F[3], Chen WH[4], Filonenko A[2]
[1]Adelphi Values, Bollington, UK, [2]Bayer Pharma AG, Berlin, Germany, [3]Adelphi Values United States, Boston, MA, USA, [4]RTI Health Solutions, Research Triangle Park, NC, USA

OBJECTIVES: As a chronic condition characterized by pelvic pain, dysmenorrhea and dyspareunia, endometriosis is associated with significant individual and societal burden. Evaluating the efficacy of endometriosis treatments is constrained by the absence of Patient-reported Outcome (PRO) measures that meet the standards of the FDA PRO Guidance. The Endometriosis Symptom Diary (ESD) and Endometriosis Impact Scale (EIS) are being developed to meet this unmet need. The establishment of the content validity of the ESD and EIS through qualitative and quantitative studies, a mixed methods approach, is presented herein. **METHODS:** The development of the ESD and EIS was informed by qualitative and quantitative stages to establish content validity. The qualitative stage consisted of face-to-face concept elicitation (n=45) and cognitive debriefing (n=31) interviews. The quantitative stage utilized Rasch Model analyses to evaluate item performance based on baseline data obtained during a prospective, observational validation study (n=268). **RESULTS:** The ESD is an electronic diary assessing patient experience of symptoms of endometriosis with